DOCUMENT RESUME

ED 178 563 TM 009 696

AUTHOR Danielson, Gary; Haupt, Edward J.
TITLE Errors on Unit Quizzes Predict Midterm Scores in a PSI Course.
PUB DATE [77]
NOTE 22p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; *Achievement Tests; Correlation; *Course Content; Criterion Referenced Tests; Higher Education; *Individual Differences; Individualized Programs; *Mastery Learning; *Predictive Ability (Testing); Scores; Undergraduate Students
IDENTIFIERS Bloom (Benjamin S); *Personalized System of Instruction

ABSTRACT

Bloom has looked to mastery learning, exemplified by the Personalized System of Instruction (PSI), to reduce individual differences in posttest performance, such as final grades and midterm examination scores. While PSI courses have raised overall performance, there has been little examination of individual differences in these courses. Data from an undergraduate PSI course in educational psychology were analyzed for the relation of errors on unit quizzes to midterm examinations and final performance. The midterms and finals were constructed of the same questions used in unit quizzes, but not all questions had been seen by all students. Errors on unit quizzes were most closely related to midterm and final exams. Bloom's model was only supported when single units or the number of attempts at mastery were studied. The model was not supported when total errors leading to mastery were used. Several scatter plots showed that a critical percentage of unit quiz errors led to sharply increasing dropoffs of midterm scores. Questions which were new to students were more affected by total error level. Bloom's model applies best to students who make few errors in achieving mastery. More effort needs to be directed to the students who make many errors. (Author/MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED178563

Errors on Unit Quizzes Predict Midterm Scores in a PSI Course

Gary Danielson

Centenary College

Edward J. Haupt

Montclair State College

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Gary Danielson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM009 696

Abstract

Bloom has looked to mastery learning to reduce individual differences in posttest performance. While PSI courses raise overall performance there has been little examination of individual differences in these courses. Data from a PSI course in educational psychology were analyzed for the relation of errors on unit quizzes to midterm and final performance. These midterms and finals were constructed of the same questions used in unit quizzes, but not all questions had been seen by all students. Errors on unit quizzes was found to be most closely related to midterm and final exams. Bloom's model was only supported when single units or the number of attempts at mastery was studied. The model was not supported when total errors leading to mastery were used. Several scatter plots showed that a critical percentage of unit quiz errors led to sharply increasing dropoffs of midterm scores. Questions which were new to students were more affected by total error level. Bloom's model applies best to students who make few errors in achieving mastery. More effort needs to be directed to students who make many errors.

Errors on Unit Quizzes Predict Midterm Scores in a PSI Course

Bloom (1971,1974) has looked to mastery learning, as exemplified in the Personalized System of Instruction (PSI) to reduce differences in performance on a variety of posttests, such as final grades, achievement tests, and midterms. Keller, too, (1968) described decreases in variability on final test performance in his first paper on PSI. While PSI courses often produce increases in overall level of performance (Hursh, 1976), there has been relatively little examination of individual differences in posttest performances and their relation to errors, study duration, and other measures of variations of effort during a PSI course.

Informally, it has appeared to us that students who take more time and/or make more errors in the process of meeting a mastery criterion often remember less well and understand less completely the material taught. Even so, Fernald, Chiseri, Lawson, Scroggs, & Riddell (1975) found no differences in exam performance between students who did or did not pass a 100% mastery criterion. The repeated requirement to take alternative forms of tests associated with higher error rates may not completely eliminate individual differences. Our belief contradicts Whitehurst and Madigan's (1975) finding that students who took more unit quizzes did as well on major exams as those who took fewer. These results of Whitehurst and Madigan may be an artifact of different withdrawal

rates, since Spencer Glick, and Semb (1977) have found that students who withdraw from PSI courses perform more poorly on major e exams and take more retests.

Questions for mastery can be written at many levels. While different patterns of individual differences in memory might be found for different levels of questions, no data exists to describe such differences. The questions available for the book used in the present study were exclusively memory-level questions. This exclusive use of memory questions provides a baseline for comparison of memory and higher-level questions. In order to determine the relation of individual differences in meeting mastery criteria and exam performance, data from a one-semester course in Educational Psychology was examined.

Method

Subjects

Data from 108 students who were given grades in an undergraduate Educational Psychology course at Montclair State College were used. These students had all enrolled for one large section of the course given by the Psychology Department. Eighty-nine students were in majors that led to teacher certification, such as Physical Education, Home Economics, Communication Sciences and Disorders, and various secondary education subject specialties. Five students were psychology majors taking the course as an elective. Fourteen other students were enrolled as uncommitted or not in majors which required Educational Psychology.

Procedure

The Educational Psychology course was adapted to use the main components of PSI (unit quizzes, mastery criteria, feedback by proctors, and optional lectures) within the time and funding constraints available at a state college.

Twelve chapters from a widely used, behaviorally based educational psychology textbook (Janice Gibson's Psychology for the Classroom) were used as the basic units for the course. Each chapter consisted of 20-35 pages of text centered around a topical heading (Memory, Associationistic Learning Theories, Disadvantaged, etc.)

On the first day of class, a 10-page outline of the course procedures, quiz schedules, office hours, and other materials were

handed out and explained. Subsequently, a quiz was given on the contents of the course outline. The quiz was the first of the 13 unit quizzes (outline plus 12 chapters). The quiz on the course outline could be taken once each day for the first three weeks of the course. The second unit, on memory, was introduced during the second week of the course, and was available during the second through fourth weeks of the course. Successive units were introduced each week.

For the three-week period during which quizzes were offered for any unit, 15 different versions of quizzes were used. The various forms of the quiz were used to reduce opportunities to cheat based on memorization of a list of answers. The quizzes were constructed from 48 questions from the instructor's manual and 30 questions from the students' study guide. The 30 questions from the study guide were formed into 15 pairs of items, one pair to be used on each quiz form. The 48 questions from the instructor's manual were formed into 6 groups of 8 questions each. The questions from each of these six groups were rearranged in a stratified manner so that nine additional sets of 8 questions were created. None of these versions contained more than two questions from any one of the six original groups of 8 questions.

Each student could take quizzes on any unit three times. A student passed a unit when no more than one error was made on a quiz. If all 10 questions on a unit quiz were correct the first time the student took the unit, the student received a "bonus"

card. The bonus card could be used in lieu of one correct answer on any unit quiz for which the student had 8 answers correct.

If a student did not answer all questions correctly, the student could receive feedback from graduate or undergraduate proctors. Feedback consisted of verbally questioning the student about the reasons for choosing an answer that was marked wrong. If it was evident that the student understood the concept for the questions but was confused by wording or other elements of the question, the student was given credit for the question. Conversely, if the student was confused about the underlying concept, the proctor explained the concept and its relationship to the question.

In addition to quizzes on units, students could also gain points toward their final grade from a midterm, final, project, and class (lecture) attendance. The midterm and final each consisted of 50 questions that had previously been used in quizzes on each unit. Unavoidably, each student saw some questions that had previously been on unit quizzes. If a student answered 90% of the questions correctly, credit was given which was equal to three unit quizzes; 80% correct produced credit equal to two unit quizzes; and 70% produced credit equal to one unit quiz. The midterm and final were both given on regular class days. Each summary test was scored immediately and the student was told how many unit quiz equivalents he had achieved, but no other feedback was given.

Each student could complete a project. The projects were activities taken from Packard's Educational Psychology (1972). Successful completion of a project resulted in the equivalent of two unit quizzes. One further unit quiz equivalent could be obtained from attendance at class lectures. Attendance was taken during 6 of the class days of the course. If the student was present on 5 of these, the unit-quiz equivalent was given.

Lecture periods (two per week) were used for one lecture on the chapter that was introduced that week and a film related to the chapter. Each graduate and undergraduate proctor was assigned a group of students who met to discuss the project, and do group or individual tutoring as necessary.

Grades for the course were determined by the number of units passed and by their equivalents in the other activities described. The grade criteria made it difficult to get an A unless a student passed all units and received an equivalent of 80% correct on both the midterm and final.

Results

Data which included the particular quizzes taken for each unit, the number of errors on each quiz, and individual item scores on the midterm and final were keypunched. The number of quizzes taken for each unit, the total errors for each unit, and the ratio of errors to quizzes taken were calculated for each unit. The midterm and final scores, and the total correct on mid- and final for the questions from each unit were also calculated.

Midterm

For each of the six units covered by the midterm, the correlation of the number correct from the midterm with the total errors per unit quiz produced consistently non-significant and negative correlations (see Table 1). These correlations include only one

---

Insert Table 1 about here

(from the unit on Instruction) which was significant, and this correlation was positive ($r(102) = .201$, $p < .05$). Students/taking were this unit immediately before the midterm; this information was more immediately available for recall. It is also possible that more errrors represent more attempts to pass the unit and thus more study. In all, these findings confirm Bloom's expectation that use of mastery procedures (with the opportunity for repeated testing) do not affect final levels of performance.

Forty-seven of the 50 midterm questions had been used on one or more of the unit quizzes. Thus, some of the questions had been seen by each student, although the number and specific questions varied from student to student. Questions previously seen by each student were determined by comparing the questions which occurred in each form of quiz with the forms which each student had completed. The mean number of previously seen questions was 7.75 (15.5% of the midterm) with a range of one to 17 questions.

Individual percentage correct scores were calculated for questions previously seen or not previously seen. Those questions which had appeared before for a student were more likely to be answered correctly (mean percent correct was 76.9) while those questions not previously seen were less likely to be answered correctly (mean percent correct was 63.2). This difference is substantially less than the 20-30% difference found by Semb, Hopkins, and Hursh (1973) for study questions included on unit quizzes.

Table 2 shows The proportion of previously seen questions answered correctly yielded a near-zero correlation with total errors on unit quizzes ($r(102) = -.67$, $p < .5$) while the proportion of questions not previously seen correlated with total errors on unit quizzes much like the midterm total score ($r(102) = -.342$, $p < .0004$).

The scatter plot for the relation of percent errors on unit quizzes to midterm total score was prepared. The scatter plot

(see Figure 1) demonstrates that midterm performance on previously

Insert Fig. 1 about here

seen questions declines rapidly once the mean error rate per quiz passes 30%. Figure 2 shows that the same relationship is apparent

Insert Fig. 2 about here

for new (not previously seen questions. However, with new questions, the critical error rate is about 20%. The total midterm produces results most like the new questions.

Final

The final was a summary exam covering the second six units from the text. Analysis of performance on unit quizzes and performance on the final was conducted like that for the midterm. Table 2 shows the correlations between unit quiz performance and

Insert Table 2 about here

performance on those same areas on the final. The italicized (diagonal) correlations show each unit and its corresponding part on the final. This matrix is much the same as the correlations for the midterm (non-significant and negative)except for one unit --evaluation--which was positively correlated with the component for that unit from the final. This unit was the one which many students were taking at the time of the final. Here again, either recent study or more attempts could account for these atypical correlations.

Next, correlations were found for the final and variables summed over all units included in the final. All these correlations were negative; the final and and the total of errors per quiz yielded an r of -.289 (df = 71, $p < .02$); the final and the total number of errors was also signfiicant ($r(71) = 261$, $p < .02$); the final and the number of quizzes taken was non-significant, but in the same direction ($r(71) = -.185$, $p < .5$). This pattern was similar to that from the midterm. Again, the variable that best predicted final performance was based on the number of errors in meeting the mastery criterion.

Figure 3 shows a plot of the percentage correct on the final

Insert Fig. 3 about here

---

as a function of the proportion of errors on unit quizzes covered by the final. The pattern of a steep dropoff after a critical number of errors occurs again. However the critical percentage of unit quiz errors was nearer 12%, instead of the 20% on the midterm. There are several possible explanations of this lower critical percentage. Better quiz passing skills may have reduced the knowledge gained; easier material may have made answering questions on unit quizzes less demanding of study; or the changed critical value may just reflect the different group (only 75 out of 108) who took the final.

Discussion

Attempts to investigate individual differences in PSI courses which only focus on small units of material, which only use questions like those used in mastery testing, or which focus solely on number of attempts to pass a unit appear to show few individual differences and support Bloom's (1971) model. Questions and tests which are novel or which are integrated over large amounts of material do not. When each unit in a mastery course was tested by itself, there were no individual differences related to errors or number of attempts to pass quizzes in mastery learning. When old questions were used. on which each student received individual review if he made an error, there were also no differences related to errors. However, individual differences in midterm performance we were found when questions which covered different aspects of the units were related to errors on unit quizzes. Thus it appears that the mastery criterion implemented in this PSI course left a residue of individual differences which were not directly tested in the unit quizzes. Students may well have been able to use a more surface form of recognition with these memory-level questions which they had seen before (Anderson & Biddle, 1975), but appear to require higher-level cognitive processes for unreviewed material which may be related to individual differences in verbal aptitude.

Students who make more than 20% errors appear to need some special consideration within PSI courses (such as special tutor-

ing or different texts) so that their retention of mastered material will be as great as those who make fewer errors while achieving mastery.

Thus the Bloom model holds better for students who make fewer errors in meeting the mastery criterion, rgardless of the number of attempts. For those students who exceed the critical error percentage, our expectations for retention and other future performances are very different, even though these students achieve mastery, too. Early identification of these students and implementation of remedial procedures may ensure that Bloom's model works for these students, too.

Since it is errors, and not their derivative, the number of attempts to pass a unit, that predicts poor retention (on a midterm), efforts should be directed to reducing errors in learning and test-taking. Few efforts other than those of programed instruction have attempted to reduce the errors that adults show in learning. Heckler, Pennypacker, and Fuqua (1975) have, however used fading (errorless learning) procedures to reduce errors and improve retention in learning facts for rote recall on true-false questions. Adaptation of such methods may be necessary to improve performance of students who do poorly on major exams in PSI courses. Only with such altered methods can Bloom's goal of reduction of individual differences be more closely approximated.

References

Anderson, R. C., & Biddle, W. B. On asking people questions about what they are reading. In G. H. Bower (Ed.), The psychology of learning and motivation (Vol. 9). New York: Academic Press, 1975.

Bloom, B. S. Mastery learning. In J. H. Block (Ed.), Mastery learning: Theory and practice. New York: Holt, Rinehart, & Winston, 1971.

Bloom, B. S. An introduction to mastery learning theory. In J. H. Block (Ed.) Schools, society, and mastery learning. New York: Holt, Rinehart, & Winston, 1974.

Fernald, P. S., Chiseri, M. J., Lawson, D. W., Scroggs, G. F., & Riddell, J. C. systematic manipulation of student pacing, the perfection requirement, and contact with a teaching assistant in an introductory psychology course. Teaching of Psychology, 1975, 2, 147-151.

Heckler, J. B., Fuqua, R. W., & Pennypacker, H. S. Errorless differentiation of academic responses by college students. Teaching of Psychology, 1975, 2, 103-107.

Hursh, D. E. Personalized systems of instruction: What do the data indicate. Journal of Personalized Instruction, 1976, 1. 91-105.

Keller, F. S. "Good-bye,teacher..." Journal of Applied Behavior Analysis, 1968, 1, 79-89.

Packard, R. W. Psychology of learning and instruction. Columbus, Ohio: Merrill, 1972.

Semb, G., Hopkins, B. L., & Hursh, D. E. The effects of study questions and grades on student test performance in a college class. Journal of Applied Behavior Analysis, 1973, 6, 631-643.

Spencer, R. E., Glick, D. M. & Semb, G. An analysis of student withdrawals in external and internal proctored personalized instruction courses. Paper presented at Fourth National Conference for Personalized Instruction, San Francisco, April, 1977.

Whitehurst, C. & Madigan, J. Slow learners in PSI courses: Do they learn less? Journal of Higher Education, 1975, 46, 55-62.

Table 1

Intercorrelations Between Errors on Unit Quizzes and Correct Responses on Representative Samples of those Units on the Midterm Examination

Midterm Units	Quiz Units					
	Memory	Behavior	Cognition	Motivation	Learning	Instruction
Memory	-.115	-.112	-.022	-.096	-.150	-.192
Behavior	-.328	-.074	-.131	-.213	-.144	-.163
Cognition	-.128	-.036	-.004	+.051	-.137	+.056
Motivation	-.166	-.177	-.080	-.150	-.278	-.146
Learning	-.272	-.123	-.044	-.118	-.066	-.137
Instruction	-.207	-.179	-.046	-.174	-.247	+.201

Table 2

Intercorrelations Between Errors on Unit Quizzes and Correct Responses on Representative Samples of Those Units on the Final Examination

Final Units	Quiz Units					
	Disadvantaged	Handicapped	Beh. Probs.	Teachers	Eval.	Testing
Disadvantaged	-.044	-.147	+.032	-.112	-.122	+.000
Handicapped	-.121	-.014	-.158	-.152	-.207	-.050
Beh. Probs.	-.204	-.050	-.061	+.021	-.156	-.075
Teachers	+.008	-.034	-.131	-.016	-.006	-.171
Evaluation	+.197	+.136	+.179	+.025	+.159	+.119
Testing	-.138	-.235	-.150	-.029	-.113	-.024

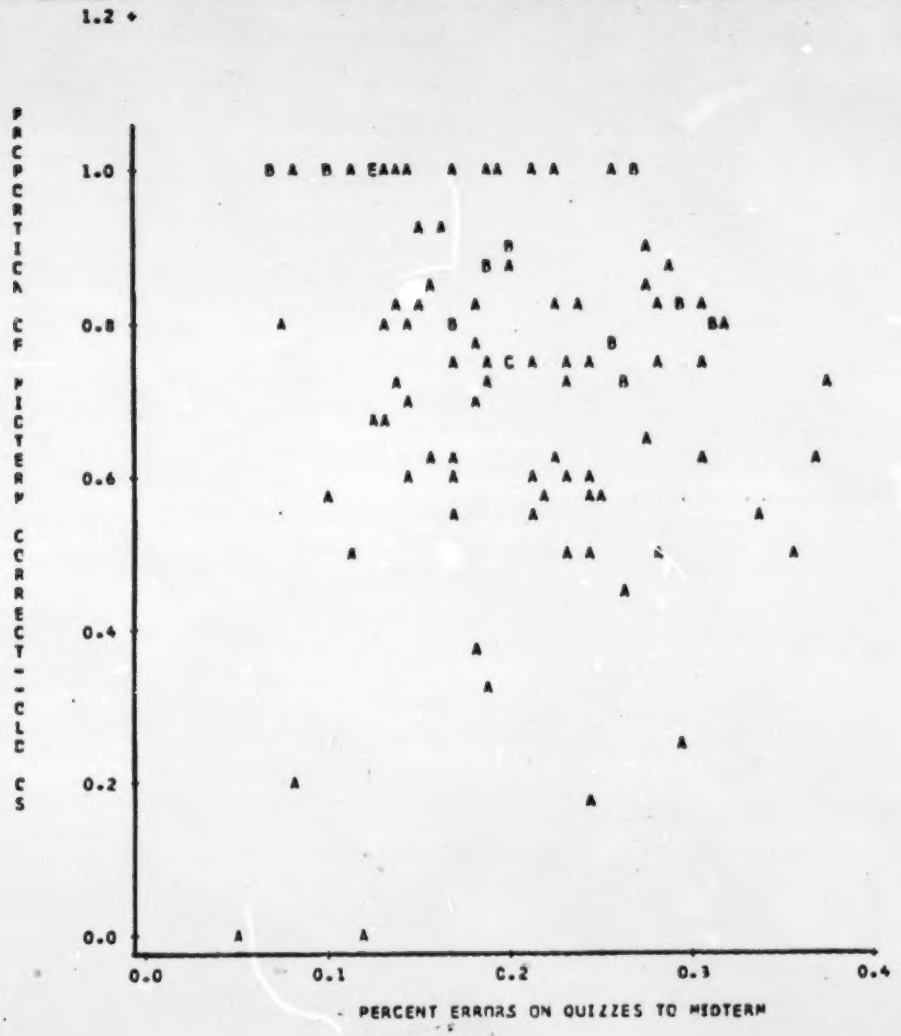

Figure 1. Percent errors on quizzes included on the midterm examination plotted against the proportion of previously seen questions answered correctly on the midterm.

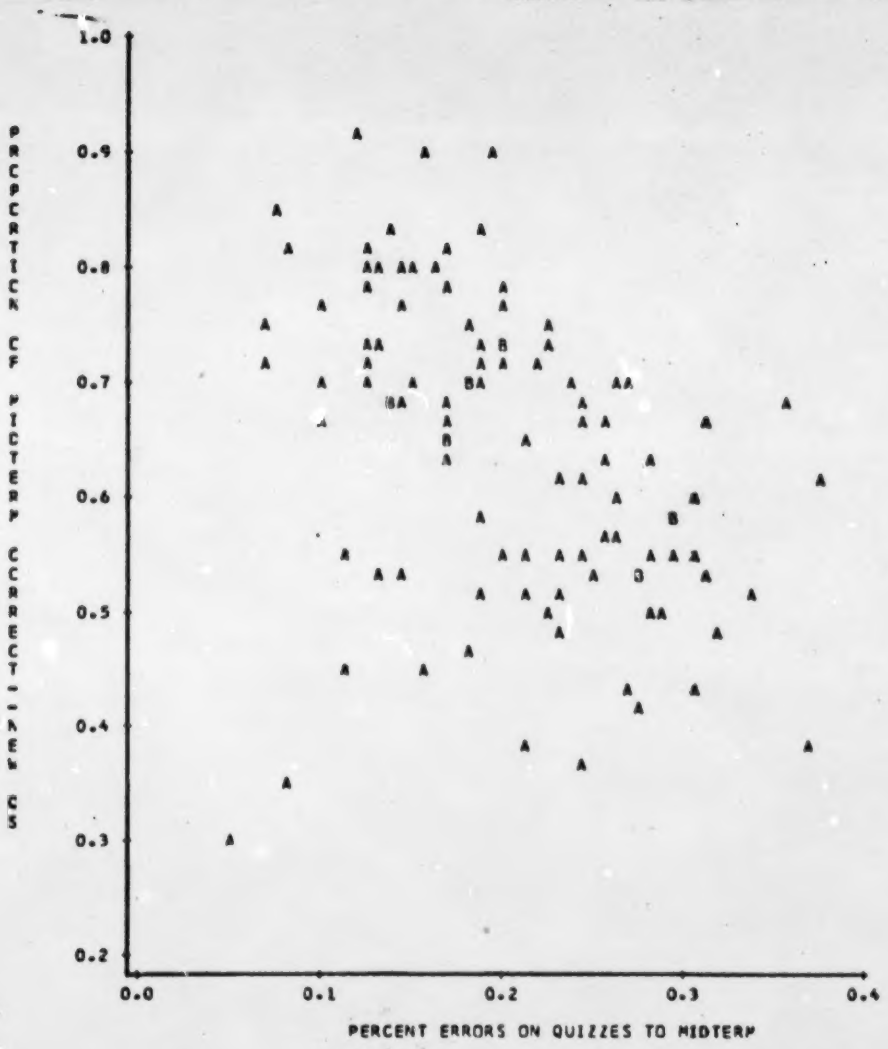

Figure 2. Percent errors on quizzes included on the midterm examination plotted against the proportion of new questions answered correctly on the midterm.

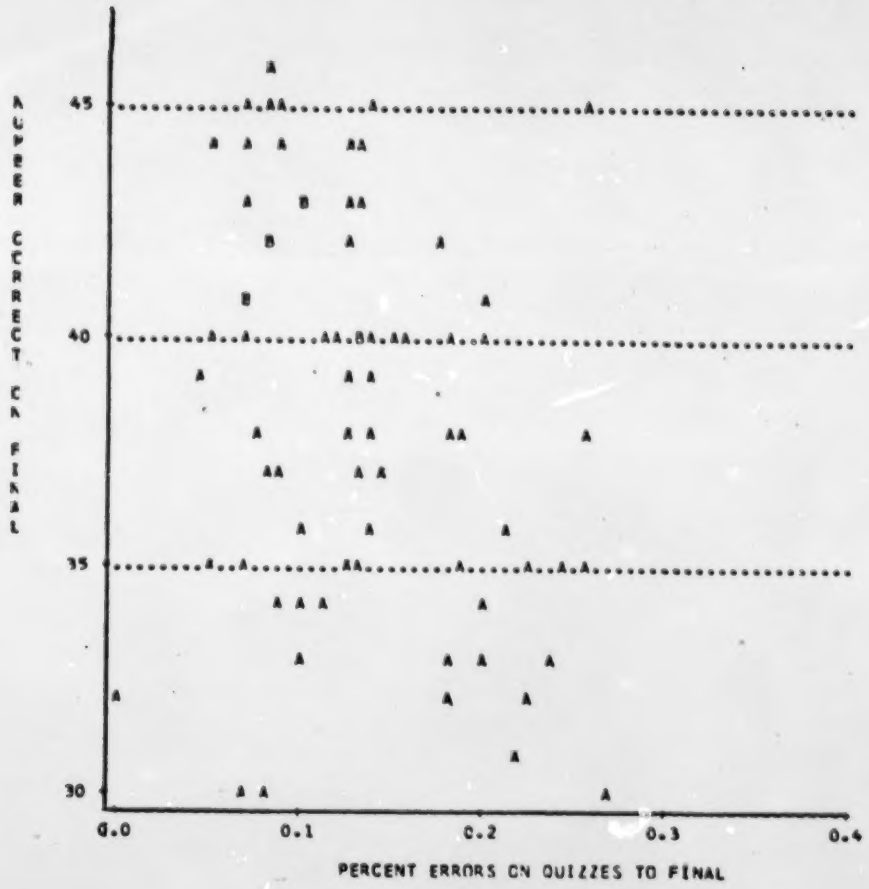

Figure 3. Percent errors on quizzes included in the final examination plotted against the number of correct responses on the final.